



## Expert and construct validity of the Simbionix GI Mentor II endoscopy simulator for colonoscopy

Arjun D. Koch,<sup>1</sup> Sonja N. Buzink,<sup>2,3</sup> Jeroen Heemskerk,<sup>3</sup> Sanne M. B. I. Botden,<sup>3</sup> Roeland Veenendaal,<sup>4</sup> Jack J. Jakimowicz,<sup>3</sup> Erik J. Schoon<sup>1</sup>

<sup>1</sup> Department of Gastroenterology, Catharina Hospital Eindhoven, Michelangelolaan 2, 5623 EJ Eindhoven, The Netherlands

<sup>2</sup> Faculty of Industrial Design Engineering, Delft University of Technology, Landbergstraat 15, 2628 CE Delft, The Netherlands

<sup>3</sup> Department of Surgery, Catharina Hospital Eindhoven, Michelangelolaan 2, 5623 EJ Eindhoven, The Netherlands

<sup>4</sup> Department of Gastroenterology, Leiden University Medical Center, Albinusdreef 2, 2333 ZA Leiden, The Netherlands

Received: 15 January 2007/Accepted: 29 January 2007/Online publication: 22 May 2007

### Abstract

**Objectives:** The main objectives of this study were to establish expert validity (a convincing realistic representation of colonoscopy according to experts) and construct validity (the ability to discriminate between different levels of expertise) of the Simbionix GI Mentor II virtual reality (VR) simulator for colonoscopy tasks, and to assess the didactic value of the simulator, as judged by experts.

**Methods:** Four groups were selected to perform one hand–eye coordination task (EndoBubble level 1) and two virtual colonoscopy simulations on the simulator; the levels were: novices (no endoscopy experience), intermediate experienced (<200 colonoscopies performed before), experienced (200–1,000 colonoscopies performed before), and experts (>1,000 colonoscopies performed before). All participants filled out a questionnaire about previous experience in flexible endoscopy and appreciation of the realism of the colonoscopy simulations. The average time to reach the cecum was defined as one of the main test parameters as well as the number of times view of the lumen was lost.

**Results:** Novices ( $N = 35$ ) reached the cecum in an average time of 29:57 (min:sec), intermediate experienced ( $N = 15$ ) in 5:45, experienced ( $N = 20$ ) in 4:19 and experts ( $N = 35$ ) in 4:56. Novices lost view of the lumen significantly more often compared to the other groups, and the EndoBubble task was also completed significantly faster with increasing experience (Kruskal Wallis Test,  $p < 0.001$ ). The group of expert endoscopists rated the colonoscopy simulation as 2.95 on a four-point scale for overall realism. Expert opinion was that the GI Mentor II simulator should be included in the training of novice endoscopists (3.51).

**Conclusion:** In this study we have demonstrated that the GI Mentor II simulator offers a convincing realistic representation of colonoscopy according to experts (expert validity) and that the simulator can discriminate between different levels of expertise (construct validity) in colonoscopy. According to experts the simulator should be implemented in the training programme of novice endoscopists.

**Key words:** Endoscopy — Simulator — Training — Colonoscopy — Validation

Training skills in endoscopy for diagnostic and therapeutic procedures is essential and requires a great deal of hands-on training [1]. Virtual reality (VR) simulators offer a promising option to train these skills extensively prior to training in real-life colonoscopy, without jeopardizing patients or causing them unnecessary discomfort [2]. The use of VR training prior to performing real flexible endoscopy on patients enables novice endoscopists to go through part of their proficiency curve before submitting patients to their relatively insufficient endoscopy skills. This might not only be advantageous for the patients undergoing endoscopy, but might also prevent complications and potential consequences resulting in medicolegal litigation. One of the simulators in the field of flexible endoscopy is the GI Mentor II (see Figure 1). VR simulators have been used extensively in different fields of expertise before applying these procedures to patients. In the United States of America simulator training is mandated by the Accreditation Council for Graduate Medical Education (ACGME) in laparoscopic procedures for surgical residents [3]. The first step is to validate the simulator construct properly and verify its didactic value, before implementing simulators in teaching programmes or developing a new curriculum for flexible endoscopy around them.



**Fig. 1.** The GI Mentor II virtual reality simulator, the setup for training in lower endoscopy.

Some studies have already been published on this subject [4–6], but the presented outcomes lacked power due to their relatively small sample sizes. In addition, some cases did not study the validity of endoscopy, but for example only the EndoBubble module, a computer simulation skills test measuring how long it takes a person to pop 20 balloons in a virtual tunnel.

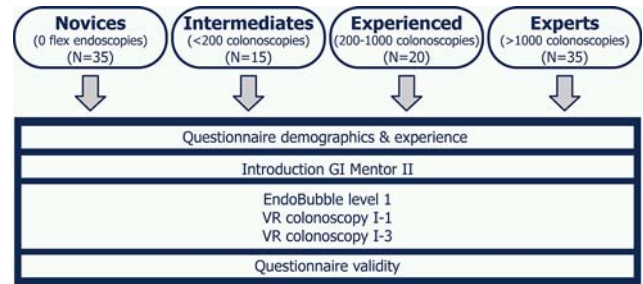
## Objective

The main objectives of this study were: (1) to establish the degree of representation of real-life colonoscopy on the Simbionix GI Mentor II VR colonoscopy simulation, as judged by experts (expert validity), (2) to determine whether the GI Mentor II simulator can distinguish between various degrees of expertise in endoscopy, judged by novice, intermediate experienced, experienced and expert endoscopists performing VR colonoscopy (construct validity), and (3) to assess the didactic value of the simulator, as judged by experts.

## Material and Methods

### Simulator

The simulator used in this study was the Simbionix GI Mentor II (Simbionix Ltd. Israel, software version 2.7.3.0) (Figure 1). The GI Mentor II can simulate upper GI tract endoscopies such as esophagogastroduodenoscopy, endoscopic retrograde cholangiopancreatographies, and endoscopic ultrasound. The lower GI tract endoscopies it simulates are sigmoidoscopy and colonoscopy. The simulator records a range of parameters upon each exercise, which can



**Fig. 2.** The study design.

be used to assess performance objectively. The endoscope used is a customized Pentax ECS-3840F endoscope.

## Participants

Participants were allocated to four groups to assess the validity and didactic value of the GI Mentor II simulator. The first group, the novices, was defined as participants without any flexible endoscopy experience; they were all medical interns or residents. The second group was intermediate experienced, with fewer than 200 colonoscopies performed before. In the third group *experienced* participants all performed more than 200 colonoscopies but fewer than 1,000. The fourth group consisted of experts, all of whom had performed more than 1,000 colonoscopies. These categories were chosen based upon several other studies, the demands for Dutch accreditation for colonoscopy, and the accreditation demands of the British Society of Gastroenterology, which advocates 200 colonoscopies under supervision during training [4, 6–8]. All persons were either invited to participate within our hospital, or participated during a national congress of the Dutch Society of Gastroenterology in spring 2006.

The groups consisted of at least 28 persons to ensure sufficient statistical power [9]. A post hoc sample size calculation based on the results for time to finish the EndoBubble task showed a minimal sample of 26 participants in the novices group to achieve a power of 0.95. Originally, the intermediate experienced and experienced participants formed one group, but as the expertise level and performance within this group varied considerably, this group was split. A schematic setup of the study design is presented in Figure 2.

## Questionnaire

All participants were asked to fill out a questionnaire on demographics and their general medical and endoscopy experience. It also included the number of endoscopies performed annually and number of years registered as a skilled professional endoscopist.

After the simulator run the participants were asked to answer questions about their appreciation of the realism of the colonoscopy exercises performed. Appreciation was expressed on a four-point Likert scale [10] varying from very unrealistic (1) to very realistic (4). Questions were asked about the realism of imaging, simulator setup, endoscope control and both haptic and visual feedback. Experts were asked whether the GI Mentor II could be used as a teaching device for novice endoscopists and whether experience on the simulator could be useful in practice.

## Simulation modules

All participants first performed the hand-eye coordination task (EndoBubble level 1) of popping all 20 balloons in the test as quickly as possible, without touching the walls. Next, the participants performed VR case numbers 1 and 3, both from colonoscopy module 1. These cases were carefully selected for their discriminative value; both cases are straightforward colonoscopies, without any abnormalities such as polyps, tumours, or inflammation. Case number 1 is a relatively easy colonoscopy to perform, whereas case number 3 is more difficult, requiring the endoscopist to apply techniques such as straightening the

endoscope during loop formation and applying torque to the endoscope shaft. The assignment given for the VR colonoscopies was to reach the cecum as quickly as possible with as little patient discomfort as possible. Patient discomfort was defined as the estimated percentage of time the virtual patient was in excessive pain and the number of times excessive local pressure was caused. Other relevant test parameters were the percentage of time spent with clear view and the number of times view of the lumen was lost. The task was considered accomplished when the cecum was reached.

### Data analysis

SPSS 13.0 software was used to perform descriptive statistics and Kruskal–Wallis tests for statistic analysis of the data. A separate analysis between groups was performed using a two-tailed Mann–Whitney exact *U* test. A *p*-value of less than 0.05 was considered significant. The data showed a nonparametric distribution, therefore the median and range of performance parameters are presented as primary values.

## Results

### Participants

Thirty-five novices, 15 intermediates, 20 experienced, and 35 expert endoscopists participated in the study. The average number of colonoscopies performed annually by experts was 445, and their mean number of years registered as a gastroenterologist was 7.7 (range 0–35 years).

### Construct validity

Data output by the simulator are presented in Tables 1 and 2. The EndoBubble task was completed faster by the experts and experienced endoscopists than by novices, with fewer wall collisions. These differences were statistically significant (Kruskal–Wallis test) (Table 1). Also the colonoscopy tasks were completed faster ( $p < 0.001$ , Kruskal–Wallis test), with less patient comfort and better visibility by experts and experienced endoscopists (Table 3). Novice endoscopists ( $N = 35$ ) reached the cecum in a mean time of 29:57 (min:sec) in colonoscopy case 3, intermediate experienced ( $N = 15$ ) in 5:45, experienced ( $N = 20$ ) in 4:19, and experts ( $N = 35$ ) in 4:56. Novices lost view of the lumen significantly more often than the other groups.

A separate analysis between groups using a Mann–Whitney exact *U* test demonstrated no significant difference between the intermediate, experienced and expert groups on all parameters. They all completed the task faster than the novices (see Table 4).

### Expert validity

The group of expert endoscopists rated the colonoscopy simulation 2.95 on a four-point Likert scale for overall realism. Anatomical representation was rated 2.58, and the simulator setup 3.14. Endoscope control scored 3.21. Haptic feedback was rated 2.57.

### Didactic value

Expert opinion was that the GI Mentor II simulator should be included in the training of novice endoscopists

**Table 1.** EndoBubble hand–eye coordination task

Experience		Time to finish (min:sec)	Number of times wall touched
Novice $N = 35$	Mean	6:56	1.9
	Median	5:58	1.0
	Range	1:24–20:25	0–20
Intermediate $N = 15$	Mean	1:56	1.1
	Median	1:41	0.0
	Range	0:54–4:02	0–5
Experienced $N = 20$	Mean	1:37	0.9
	Median	1:21	0.0
	Range	0:43–5:33	0–9
Expert $N = 35$	Mean	1:24	0.3
	Median	1:13	0.0
	Range	0:49–3:25	0–2
Kruskal–Wallis	Chi- square	63.151	9.374
	Asymp. sign	.000	0.025

(3.51 on a four-point Likert scale) and that expertise gained on the simulator was considered applicable in a clinical curriculum (rated 3.29 out of 4). The simulator was not considered suitable for certification of trained endoscopists (rated 2.29 out of 4).

## Discussion

This study represents the largest and most detailed study on the validity of this type of colonoscopy simulator so far. The data show that the simulator can discriminate clearly between endoscopists of different expertise levels performing different colonoscopy tasks. Differences were statistically significant using relatively large sample sizes in all three exercises, the EndoBubble task as well as cases number 1 and 3. The difference between our study and previous studies by others is that we focused on the basic aspects of navigation for colonoscopy itself, rather than on the hand–eye coordination task alone, used for example in the study by Ritter et al. [4], and that we included more participants in four separate groups with different levels of expertise [4–6, 11, 12]. In this way we were able to demonstrate that the GI Mentor II can distinguish between expertise levels up to the level of an intermediate experienced endoscopist, who has performed around 200 colonoscopies. In a similar study Sedlack et al. [5] describe a limited construct for a different simulator (AccuTouch, Immersion Medical). Felsner et al. [11] demonstrated differences between novices and experts in large sample sizes but did not compare novices to intermediate levels of expertise. In this study we have demonstrated convincing expert validity for colonoscopy on the GI Mentor II virtual simulator. This in contrast to other studies focusing on the EndoBubble task as a validation study [4] and not dealing with the subject of expert validity [4, 6, 7, 11, 12].

The colonoscopy tasks were considered as accomplished once the participants reached the cecum. Asking the participants to inspect the mucosa on the way back through the colon does not, in our opinion, provide a proper representation of the endoscopist's skills in manoeuvring through the colon, as other aspects besides the basic navigation skills of the endoscopist could

**Table 2.** Colonoscopy module 1, cases 1 and 3

Experience			Time to reach cecum (hour:min:sec)	% of time spent with clear view	Lost view of lumen	Excessive local pressure	% of time patient was in pain	Excessive loop formed
Case 1	Novice $N = 35$	Mean	6:47	96	0.4	0.5	13.3	0.83
		Median	6:16	97	0	0	11	0
		Range	1:53–15:08	82–99	0–3	0–3	0–44	0–6
	Intermediate $N = 15$	Mean	1:36	97	0	0	8	0.6
		Median	1:40	98	0	0	5	0
		Range	0:55–2:52	91–100	0	0	0–30	0–3
	Experienced $N = 20$	Mean	1:23	98	0	0.2	9.2	0.7
		Median	1:21	98	0	0	8	1
		Range	0:48–2:43	89–100	0	0–1	0–27	0–3
	Expert $N = 35$	Mean	1:23	98	0	0	14.5	1.49
		Median	1:17	98	0	0	12	1
		Range	0:42–3:16	94–100	0–1	0	0–57	0–10
Case 3	Novice $N = 35$	Mean	29:57	86	3.2	3.89	2.2	4.77
		Median	23:42	85	3	3	0	1
		Range	4:48–1:28:19	72–96	0–12	1–14	0–24	0–34
	Intermediate $N = 15$	Mean	5:45	89	1.1	2.1	0.9	1.13
		Median	4:21	92	1	2	0	0
		Range	2:28–13:41	78–97	0–4	0–6	0–4	0–8
	Experienced $N = 20$	Mean	4:19	91	0.6	1.9	1.0	1.6
		Median	3:50	91	0	1	0	1
		Range	2:27–7:02	73–99	0–3	0–8	0–4	0–9
	Expert $N = 35$	Mean	4:56	89	0.9	1.6	2	2.51
		Median	4:03	90	1	1		1
		Range	1:38–15:39	68–99	0–4	0–6	0–10	0–12

**Table 3.** Statistics colonoscopy module 1, cases 1 and 3

		Time to reach cecum	% of time spent with clear view	Lost view of lumen	Excessive local pressure	% of time patient was in pain	Excessive loop formed
Case 1	Chi square	69.043	13.889	18.415	19.783	7.101	10.691
	Asymp. sig.	0.000	0.003	0.000	0.000	0.069	0.014
Case 3	Chi Square	65.559	6.978	41.936	28.794	4.284	4.856
	Asymp. sig.	0.000	0.073	0.000	0.000	0.232	0.183

Kruskal-Wallis test

**Table 4.** Differences between groups module 1, cases 1 and 3

		Time to reach cecum	% of time spent with clear view	Lost view of lumen	Excessive local pressure	% of time patient was in pain	Excessive loop formed
Case 1	Novice vs. intermediate	0.000	0.177	0.039	0.013	0.070	0.743
	Intermediate vs. experienced	0.166	0.617	1.000	0.244	0.385	0.547
	Experienced vs. expert	0.962	0.621	1.000	0.043	0.077	0.020
	Intermediate vs. expert	0.141	0.259	1.000	1.000	0.018	0.009
Case 3	Novice vs. intermediate	0.000	0.104	0.000	0.004	0.584	0.040
	Intermediate vs. experienced	0.257	0.394	0.285	0.503	0.771	0.184
	Experienced vs. expert	0.969	0.297	0.153	0.942	0.154	0.726
	Intermediate vs. expert	0.326	0.757	0.870	0.416	0.111	0.090

Mann-Whitney two-tailed test, exact significance

influence the performance parameters provided by the simulator considerably in this case. This might lead to very different end times depending, for example, on the carefulness of the endoscopist.

This study demonstrates that the GI Mentor II simulator offers a convincing, realistic representation

of colonoscopy according to experts. The overall assessment was good. Expert opinion was that the simulator can be used as a teaching tool for novice endoscopists. The simulator's haptic feedback is doubtful. Inexperienced residents can be trained in the skills necessary in flexible endoscopy such as steering

control, straightening the endoscope during loop formation and applying torque up to a certain level.

## Conclusion

The current study demonstrates that the GI Mentor II simulator offers a convincing, realistic representation of colonoscopy according to experts (expert validity) and that the simulator can discriminate up to the level of intermediate experienced endoscopists (construct validity) in colonoscopy. In the cases used the simulator could not discriminate between intermediate, experienced and expert endoscopists. The next step will be a study to determine whether novice endoscopists can develop a learning curve that will actually improve their endoscopic skills applied to real patients.

## References

1. Thomas-Gibson S, Williams CB (2005) Colonoscopy training new approaches, old problems. *Gastrointest Endosc Clin N Am* 15(4): 813–827
2. Sedlack RE, Kolars JC, Alexander JA (2004) Computer training enhances patient comfort during endoscopy. *Clin Gastroenterol Hepat* 2: 348–352
3. Roberts KE, Bell RE, Duffy AJ (2006) Evolution of surgical skills training. *World J Gastroenterol* 12(20): 3219–3224
4. Ritter E, McClusky Dr, Lederman A, Gallagher A, Smith C (2003) Objective psychomotor skills assessment of experienced and novice flexible endoscopists with a virtual reality simulator. *J Gastrointest Surg* 7: 871–878
5. Sedlack RE, Kolars JC (2003) Validation of a computer-based colonoscopy simulator. *Gastrointest Endosc* 57(2): 214–218
6. Mahmood T, Darzi A (2003) A study to validate the colonoscopy simulator. *Surg Endosc* 17(10): 1583–1589
7. Datta V, Mandalia M, Mackay S, Darzi A (2002) The PreOp flexible sigmoidoscopy trainer. Validation and early evaluation of a virtual reality based system. *Surg Endosc* 16: 1459–1463
8. BSG Guidelines (1994) Gastro-Intestinal Endoscopy in General Practice. *Gut* 35: 1342
9. Field A (2005) *Discovering statistics using SPSS*. Sage Publications Inc.
10. Likert R (1932) A Technique for the Measurement of Attitudes. *Arch Psych* 140: 55
11. Felsher JJ, Olesevich M, Farres H, Rosen M, Fanning A, Dunkin BJ, Marks JM (2005) Validation of a flexible endoscopy simulator. *Am J Surg* 189(4): 497–500
12. Grantcharov TP, Carstensen L, Schulze S (2005) Objective assessment of gastrointestinal endoscopy skills using a virtual reality simulator. *JSL* 9(2): 130–133